Adriana Sá
adrianasa@clix.pt

ESAD.CR/ IPL

Atau Tanaka
a.tanaka@gold.ac.uk

Goldsmiths,
University of London

The variables of spatial presence: a parametric model

The term 'spatial presence' refers to the feeling of presence in a mediated space. This subjective experience has been discussed in media theory, sound art, film and performance. It depends on multiple variables, or parameters. This paper presents a parametric model that can be used to analyze those variables and their relationships. It exposes methods to assess interaction, characteristics of sound and image, audio-visual relationship and physical setup. It also exposes methods to assess how these variables intertwine in perceptual experience. The model draws from perception science, interaction design, music and audio-visual theory. It is applicable to the broad diversity of aesthetical options and technical platforms, facilitating the analysis of spatial presence in any performance. One can also discard part of the parameters so as to analyze installations, sound art and film.

1 INTRODUCTION: WHERE IS 'PRESENCE'?

In media theory, Draper et al. proposed that spatial presence depends on attentional processes [1998]. Schubert described it as "a feedback of unconscious processes of spatial perception that try to locate the human body in relation to its environment" [2009:15]. And Sacau et al. stressed that it is informed by the properties of media stimuli as well as by individual psychological factors, such as the capability to be immersed [2008]. These authors concluded that in radio and TV the sense of spatial presence depends more on the active suspension of disbelief than on the properties of media stimuli. Conversely, in complex interactive virtual environments designed to convey participatory interaction, the sense of spatial presence would depend more on the properties of stimuli. These conclusions might be arguable, but their starting point is solid: spatial presence depends on stimuli as well as individual predisposition. Moreover, stimuli can condition predisposition and attention to a lesser or greater degree.

Music and sound art can make us feel present in a space that is suggested through recognizable sounds; as an example, the sound of the ocean can convey the sensation of a maritime environment. However, we can equally speak of presence when sounds are not recognizable. The composer Francisco López wrote about a particular sense of spatial presence in music: "Being 'inside' the sound (instead of listening to it) creates a strong feeling of immersion where your own body moves into the perceptive background" [2004]. In sound art philosophy, Salomé Voeglin stated that audition differs from vision because "seeing happens (…) away from the seen"; according to her, this distance enables a detachment and objectivity that presents itself as truth [2010:12].

The term "detachment" must be put into perspective, because narrative film and media can forge audio-visual relationships so as to convey presence in a suggested space beyond the screen. In the physical world, we use multiple cues for detecting the material environment, including the position of objects and events. Film and media can create a convincing sense of space with much fewer cues, because each event activates unconscious memories of similar real-world events. Michel Chion coined the term *superfield* to designate the sound space created by multi-track sound and multi-speaker placement in the movie theatre [1994]. The superfield does not depend on what we see moment by moment on screen; instead it provides "a continuous and constant consciousness of all the space surrounding the dramatic action" [1994:150]. The superfield also exists with digital 3D environments, whether the output sound is emitted through a stereo system or a multichannel system. The term '3D positional audio effects' applies to software that makes sound spatialization depend on visual dynamics. This audio-visual correspondence mimics how we perceive the physical world, conveying a sense of 'immersion' in the digital world.

In electronic music performance, the mediated space of presence can be centered on the performer, or expand to the environment. Simon Emerson's distinction between local functions and field functions is useful to illustrate this idea [2007:92]. Whilst local functions seek to extend, but not

break, the perceived relation of human performer to sounding result, field functions create a context, a landscape or an environment where local activity may be found. Emmerson prefers not to play down this division, but add supplementary dimensions to it - local can become field, and vice-versa. He uses the term *performative arena* to describe the relation between performer, audience, space, sound sources and events.

Other researches reflect similar ideas. Birnbaum et al. consider "the total physical area inhabited by the instrument/ system" as a variable in performance [2005]. And Ciciliani distinguishes between *centripetal* and *centrifugal* performances [2015]. In centripetal performances, the performer is visible and the centre of attention. The relation between his physical action and the sonic results is clearly perceivable, and the sound sources are placed near to him. In centrifugal performances, the performer is in a hidden position. He functions as a controlling rather than enacting entity; there is little or no correspondence between his physical actions and the sonic results, and the sound sources are spread in space.

The notion of performative arena is very useful to our investigation: we can say that it relates the physical and psychological space of the work. It corresponds to how the work creates the potential space of presence. That might depend on the characteristics of the sound, the image and the audio-visual relationship, on the performer's interaction with the system, the speakers' placement, the spatial relation between performer and visual projection, if any, the lighting, the physical architecture, and the audience location. The feeling of presence is also informed by individual predisposition, but that does not depend on the work, and we will not attempt to parameterize experience itself. Instead, we will glean a set of verifiable variables, and provide methods to interpret their relationships.

We examined several of those variables in previous publications, which introduced methods for analyzing sonic expression and sensory dominance [Sá 2013; Sá, Caramiaux and Tanaka 2014; Sá 2017]. Those previous investigations are a stepping stone to the work presented in this paper. Section 2 will illustrate a set of low-level variables, or parameters, with artistic examples from the sonic, the visual and the audio-visual domain. Section 3 will introduce two high-level parameters, so as to characterize semantics and the performative arena. It will examine how the high-level parameters are informed by low-level parameters, and provide artistic examples as well.

2 LOW-LEVEL PARAMETERS

This section parameterizes interaction, sonic and visual dynamics, audio-visual fit, and performer's position relatively to image. Each parameter is gleaned and exemplified independently from the others.

2.1 Interaction

Birnbaum et al. [2005] and Magnusson [2010] created parametric models to analyze interaction with digital music devices. Both models include

parameters related to the performer's control over the device and the fore-knowledge required in interaction. We can summarize those variables into a single parameter: the performer's cognitive effort, including all kinds of information processing, conscious and unconscious.

In a previous publication we elaborated on how different levels of effort convey different notions of musical expression (Sá 2017). This is equally applicable to the sonic, the visual and the audio-visual domain. As a parameter, effort can be quantified as follows:

- **Little effort** means one of two things: either the work does not depend much on real-time interaction, or the relationship between deliberate human agency and output results is linear and clearly perceivable.
- **Medium effort** means that interface behaviors are complex, with a certain potential for unpredictability; the performer needs particular skills to play the instrument, but a sense of immediacy conveys fluency and timing, and/or technical configurations rule out undesired outcomes.
- **High effort** implies particular skills and/or high cognitive demand; the interaction with the system does not feel immediate, and/ or the system does not rule out any outcomes.

An example of little interaction effort can be found in Phill Niblock's *Movements of People Working*, performed since 1973.[1] These works show repetitive movements of manual labor combined with massive drones of sound, rich in harmonics and overtones. There is no technological connection between sound and image, and yet, the connection between what we hear and see is undeniable. Niblock's real-time interaction with the laptop consists of choosing media files and pressing the playback button from time to time. The visual component is created in advance, and at least part of the music is pre-recorded as well (other musicians might play live, following his minimalistic scores).

Another example of little effort can be found in *Music for Solo Performer* by Alvin Lucier (1965)[2], where he uses a brain interface to activate multiple percussion instruments. The strength of the brain wave signal is inverse to the actuation upon the instruments - the sound is loudest when he is in a kind of meditative state. This work raises an interesting issue: the interaction with an effortful interface can be effortless. Indeed, brain waves are hard to control. But in an interview Lucier explained that he didn't want to show mind control, because he preferred the discovery of how his brainwaves sounded.[3] To him, composition was about how to deploy the loudspeakers and what instruments to use.

Medium effort can manifest in that which Jeff Pressing called dynamic complexity [1987]: a rich range of behaviors over time, an adaptation to unpredictable conditions, a monitoring of results in relation to a reference source, and an anticipation of changes in oneself or the environment. We can say that medium effort implies behavioral deviations, and reactions to those deviations.

A musical example is in a performance by Joel Ryan (electronics) and Evan Parker (soprano saxophone), 2012.[4] The sounds of the saxophone and the electronics intertwine like a braid, with floating tonal centers.

1. https://www.youtube.com/watch?v=mCKtqsy9gcY; https://www.youtube.com/watch?v=LJtIZskdHOc

2. http://daily.redbullmusicacademy.com/2017/05/alvin-lucier-music-for-solo-performer

3. 1986 interview with Ev Grimes for Yale's Oral History of American Music

4. https://www.youtube.com/watch?v=IQ4DqRgtbHc

5. https://www.youtube.com/watch?v=9mg1weOHWSs

7. https://vimeo.com/54006161

They converge when their loudness and tone are the same, then they cause attention to focus on subtle tonal shifts, diverging progressively as one timbre emerges from the other, so as to converge again.

Another example of medium effort is in an audio-visual performance by Steina Vasulka at the Smithsonian American Art Museum[5] (2012). The work pertains to her *Violin Power* series, in which she controls video footage with an electric violin with MIDI output[6] (the Zeta Violin). We can speak of audio-visual tension whenever she transits between different segments of video footage; these transitions are abrupt at the visual level, but not at the sonic level. The subsequent return to audio-visual synchrony creates a convergence, which causes a sensation of release.

A high level of effort conveys yet another type of expression. A paradigmatic example can be found in the work of Martin Howse, which deals with *psychogeophysics* since 2009.[7] Howse investigates the links between geophysical phenomena, software and the human psyche, proposing a return to animism within a critical misuse of scientific technology. In performance, his interfaces combine a diversity of chemical substances, earth materials and computers. Similarly to Lucier, Howse emphasizes discovery, as opposed to control. But very differently, the output is highly dependent on his real-time decisions and actions.

2.2 Dynamics

Another relevant question in the analysis of spatial presence is how the dynamics of sound and image drive attention. Drawing from neuroscience and psychology, we created a taxonomy of continuities and discontinuities related with intensity and attention (Sá 2013). It was useful to define intensity as the 'neural impact of any change in the chain of stimuli causing an increase in neural activity'. Since neural activity reflects attention, we can quantify intensity based on how attention works.

Attention is automatic when driven by salient events, such as the sudden appearance or disappearance of a stimulus. Such events counteract biophysical expectations, causing a great increase in neural activity. Conversely, attention is under individual control when expectations are fulfilled, requiring little neural activity. It is important to note that expectations depend greatly on the panorama - previous and simultaneous events, and the duration of experience. Also, the threshold between deliberate and automatic attention can be fuzzy. That is because attention causes us to optimize perceptual resolution, so as to better process information related to the attention target [Knudsen 2007]. Deliberate attention can make the intensity of any detail changes grow exponentially. As those inform expectations, we also become more susceptible to automatic attention.

In summary, intensity is proportional to perceived discontinuity. It depends on the event itself, on the panorama, and a person's perceptual resolution. We created the following taxonomy of continuities and discontinuities related with intensity and attention:

- **Steady continuity** has no intrinsic motion; it is of lowest intensity, dispensing with attention. Attention is likely to deviate and focus upon any simultaneous stimuli, or upon internal states.
- **Progressive continuity** occurs when successive, non-abrupt events display a similar interval of motion. It fulfils the expectation that once something begins to move in a certain direction, it will continue to move in that direction (Gestalt of *good continuation*).
- **Ambivalent discontinuity** refers to the threshold between continity and discontinuity. At low perceptual resolution – when a person is not paying attention - the foreseeable logic is shifted without disruption. At high resolution, discontinuities become more intense – even disruptive. Higher intensity implies greater attention/ neural activity, and lower intensity implies lesser attention/ neural activity.
- **Radical discontinuity** violates psychophysical expectations, prompting automatic attention. It is of highest intensity, given the great demand of cognitive processing. Radical discontinuities are always prioritized in the stimuli competition to reach conscious awareness.
- **Endogenous continuity** corresponds to the mental representation of perceptual motion; it can combine all the other types of continuities and discontinuities. We use the term endogenous so as to stress that perceiving a coherent relationship between them depends greatly on the individual.

We can say that every artistic work aims the audience to experience endogenous continuity, but all the other types of continuities and discontinuities can be illustrated with paradigmatic examples.

A musical example of steady continuity is Elaine Radigue's *Triologie de la Mort* (1998),[8] a three-hour drone piece, where we hardly perceive any overtones/ harmonic variation. The work follows the path of the continuum of the six states of consciousness according to Tibetan Buddhism. An audio-visual example of steady continuity is La Monte Young and Marian Zazeela's *Dream House*,[9] an ongoing installation (since 1962). The work defines a vibratory space through the combination of continuous sound frequencies and continuous light frequencies, experimenting on how people are drawn to inhabited it.

The notion of progressive continuity can be illustrated with any gradual increase or decrease in loudness, tonality, brightness, color, density, rhythm or time length. A concrete example can be found in Gary Hill's film *Black and White Text* (1980),[10] which explores a relationship between geometric black and white figures and human voice. As the work unfolds, the intervals between the words and the visual changes become progressively shorter, while sound layers accumulate and rectangles multiply on screen. Importantly, we cannot speak of progressive continuity unless we perceive motion. If the progression happens so slowly that we cannot apprehend any change (as happens in Radigue's *Triologie*), we should rather speak of steady continuity.

Whereas progressive continuity entails motion in a clearly perceivable direction, ambivalent discontinuity entails multi-directional motion.

8. CD released by Experimental Intermedia Label

9. https://www.youtube.com/ watch?v=3ahgq-zVQLc

10. https://www.youtube.com/ watch?v=bg1O3NcPwBg

11. https://www.youtube.com/watch?v=ojVX8FWYc4g

12. https://expcinema.org/site/en/dvd/phil-niblock-movement-people-working

13. http://www.fondation-langlois.org/html/e/page.php?NumPage=481

14. CD released by Blast First Petite, https://www.youtube.com/watch?v=kPhkH1da08o

15. https://vimeo.com/49873167

A paradigmatic example is in Thomas Wilfred's *Lumia* (1930s).[11] The Lumia were dynamic light-paintings performed with the *Clavilux*, a visual instrument that made use of multiple projectors, reflectors and colored slides. These works consist of polymorphous, fluid streams of color, which invite attention to focus on subtle detail changes. As another example, Niblock's Movements of *People Working*[12] invites perception to focus on the nuances of sound and image - the repetitive movements of manual labor and the continuous mass of sound. As we increase perceptual resolution, the detail variations become more intense.

The fuzzy threshold between ambivalent discontinuity and radical discontinuity can be exemplified with *Stellar* (1993), one of Stan Brakhage's silent abstract films. Inspired by Norman McLaren, he drew directly on film. The film has elements of continuity; all frames are made with same technique and the same colors. And yet, as we focus on the visual detail changes proliferating in the fast-changing frames, we also become very sensitive to how the interplay of discontinuities grounds the construction of time.

Radical discontinuities can be used so as to create rhythmic patterns, as happens in Vasulka's *Heraldic View* (1974),[13] or in club music. When the duration of the experience is short, each abrupt event prompts automatic attention, causing a sudden increase in neural activity. Yet after a while, the sequence of elements fulfils expectations, as happens with any pattern; in this way, radical discontinuity turns into steady continuity. Alternatively, radical discontinuities can be explored so as to tease and counterpoint expectations. For example, the *Corona* CD by Pan Sonic (2010)[14] uses radical discontinuities, but sparsely. In the audio-visual domain, Ryoji Ikeda's performance *Superimposition* (2012)[15] creates radical discontinuities with sudden blackouts; the contrast with moments of progressive and steady continuity makes those discontinuities more intense.

2.3 Audio-visual fit

The way perception prioritizes sensory information is influenced by the dynamics of sound and image, but the surplus of meaning produced by the audio-visual relationship is equally influential. In audio-visual theory, Chion coined the term *added value* to describe the surplus of synchronization [1994]. It is important to not misinterpret the term, because the meaning of the audio-visual composite is not really added to the meanings of the sound and the image. On the contrary, it tends to override those meanings.

In experimental psychology, Kubovy & Schutz showed that the aural discounts the visual and the visual discounts the aural based on concepts of causation [2009, 2010]. They coined the term *ecological fit* to describe how automatic interactions between the senses draw from those concepts. They demonstrated this with a study about the perception of a percussive action, recorded on video. The study involved the image of a marimba player, a marimba sound and a piano sound. When sound and image were desynchronized, the perceived marimba sound was shortened so as to coincide with the visible impact, but the piano sound was perceived in full

length. Moreover, the marimba sound was shortened when preceded by the visible impact, but not when the stimuli sequence was reversed.

The greater is the ecological fit, the more we ignore any diverging sensory information. We can explain this in terms of "cognitive efficiency". A high level of fit leads to integrated perceptual encodings and representations, which require less neural activity than separated encodings and representations [Brown and Boltz 2002, Boltz 2004]. Drawing from perception science and audio-visual theory, we conducted a study on perceived audio-visual relationships [Sá, Caramiaux and Tanaka 2014]. It led us to distinguish three levels of ecological fit:

- **High fit** means that the audio-visual relationship conveys conclusions about causes and effects. Perception prioritizes information that converges with those conclusions, producing integrated mental representations. High fit is of low intensity, because it requires little cognitive processing.
- **Medium fit** means that one senses causation without understanding the base cause and effect relationships. It is of medium intensity and requires a medium level of cognitive processing. It conveys perceptual chunking, but the process of audio-visual binding remains ambivalent: one can form integrated as well as separated representations of the sounds and the images.
- **Low fit** means that perceptual binding is weak because the sound/image pairing does not activate previous memories of causation. Discerning an audio-visual relationship requires perception to create new chunks of memory, with a large amount of cognitive processing. That means high intensity.

High ecological fit can be illustrated with Norman McLaren's abstract animation film *Dots* (1940),[16] where sound and image are synchronized one-to-one. The visual elements consist of dots, which McLaren painted directly onto clear frames of film. The sounds were created in the same way, with dots painted directly into the area on the filmstrip usually reserved for the soundtrack. Another good example is *Noise Fields* (1974),[17] a fully synchronized video by Steina and Woody Vasulka. Made with video processor machines, this work visualizes and sonifies the deflected energy of the electronic signal.

Beyond film and video art, many systems and instruments were designed to emphasize the union of audition and sight through one-to-one synchronization. The *Ocular Harpsichord* created by Louis Bertrand Castel (1730) is an early example. It consisted of a harpsichord with colored glasses and curtains; when a key was struck, a corresponding curtain would lift briefly to show a flash of corresponding color [Moritz 1997]. As a contemporary example, creative practitioners working with 3D positional audio-effects are usually keen on how these technologies create a high audio-visual fit. Tarik Barri's *Versum* (developed since 2008)[18] is an example. The digital 3D environment is seen and heard from the viewpoint of a moving virtual camera. Controlled in real-time by means of a joystick this camera moves through space, similar to how first person shooter games work. Within this space,

16. http://vimeo.com/15919138

17. http://www.fondation-langlois.org/html/e/page.php?NumPage=483

18. http://tarikbarri.nl/projects/versum

the author places objects that can be both seen and heard, and like in reality, the closer the camera is to them, the louder one hears them.

The perceptual effects of medium ecological fit were demonstrated in our study about audio-visual relationships, mentioned above. An audio-visual mapping exhibiting medium fit is what we call a *fungible mapping.* It combines synchronized and non-synchronized components, exhibiting complexity enough to be confusing. In our study, the participants were aware of a causal relationship, and aware of not distinguishing the base cause and effect relationships. As they could not segregate converging and diverging information, their sense of causation extended to the mapping as a whole.

It is worth mentioning that this study was greatly motivated by the development of an audio-visual instrument,[19] which combines an acoustic string instrument and software that uses a game engine with 3D positional audio-effects [Sá 2013, 2014, 2017]. The study clarified how we could confound the cause-effect relationships despite the technical platform, which is intended to maximize the audio-visual fit. Also, in performance the relation between physical gesture and instrument output is sometimes synchronized and other times not. And we use two stereo systems crossed in space, which blurs the relation between the visible sound emitters on the screen and the corresponding sounds emitted through the loudspeakers.

Another way of creating medium fit can be found in the performances of *Sensors Sonics Sights*[20] (SSS) [Tanaka 2014]. The trio uses ultrasound sensors to modulate 3D imagery, while the sound is produced with a Theremin and the Biomuse, an instrument that operates based on EMG biosignals. The audience is likely to sense a causal connection between the performers' gestures and the sonic/ visual outputs - it truly exists. Nevertheless, the nature of the instruments confounds the base cause and effect relationships. There is no technological connection between sound and image, but points of sensory unison convey perceptual binding. As perception doesn't segregate the elements that produce a sense of causation from the elements that do not, the feeling of causation extends to the audio-visual relationship as a whole.

Yet a different way of creating medium fit can be found in many live coding performances. Live coding is a performance practice where software that generates music and/or visuals is written and manipulated as part of the performance [Collins et al. 2003]. Usually, the code is projected on a screen so that people can see the process. Yet, the cause and effect relationships are often confounding. Alex McLean sometimes purposefully obscures his code to make it more difficult to read, while still showing some of the activity of the edits [Sá et al. 2015:27]. Another example is Thor Magnusson's *Threnoscope*, a live coding system for microtonal drone music [2014]. The digital cause-effect relationships are exposed with a graphic notation system and real-time programming code. Yet, even audience members who know the system and the programming language won't fully understand the cause-effect relationships, because the visible code is relatively high-level, and the system is complex. In other words, medium fit is also compatible with consistent synchrony.

19. http://adrianasa.planetaclix.pt/research/practiceOverview.htm

20. https://babiole.net/en/sss-sensors-sonics-sights/

21. https://www.youtube.com/watch?v=GPhAvyrZ28o & http://www.ryojiikeda.com/project/testpattern/

22. https://www.youtube.com/watch?v=yY5VryfCRig

23. http://www.medienkunstnetz.de/themes/overview_of_media_art/perception/5/

It is easy to find examples of low fit between physical gesture and system output, particularly in laptop performances such as Ikeda's *Test Pattern* (2008).[21] But when we watch an audio-visual performance or a film we are driven to perceive – and imagine - connections between the sounds and images, even when their fit is low and perceptual binding is weak. Indeed, one can extrapolate meanings from video images of a summer beach coupled with sounds of a warzone, even if there is no synchrony. In any case, Tony Conrad's *The Flicker* (1966)[22] is perhaps an example of low fit in film. While the sound is steady and continuous, the visual changes are abrupt and disruptive. The image consists of only five different frames: a warning frame, two title frames, a black frame and a white frame. Light and dark frames alternate with a rhythm that has been assessed to produce after-images, seeing spots, and similar phenomena.[23]

2.4 Performer position relative to a moving image

We can use an additional low-level parameter to characterize performances with moving images - one that summarizes how the performer's physical body relates to the image. That can indicate whether the visual focus is the image itself, or the relation performer-image. The following three types of arrangement are easily distinguishable, dispensing with illustrative examples:

- **Integrated** means that the image and the performer's physical body form a single visual scene, as happens when an image is projected upon a performer.
- **Separated** means that the image is separated from the performer, who is nevertheless visible. This type of arrangement can divide attention, or deviate attention from the performer.
- **Hidden** means that the performer is not visible. The audience does not see their agency, but knowing what type of interface is being used can influence how the work is perceived.

The way the physical setup influences attention also depends on speaker placement, lighting, audience distribution and physical architecture. Every detail counts in the audience's experience. But a parametric model should be easy to use, and we do not need to parameterize each variable independently. Instead, we can use high-level parameters to provide cues about variables that our low-level parameters do not address. Fig. 1 shows the low-level parameters from our model.
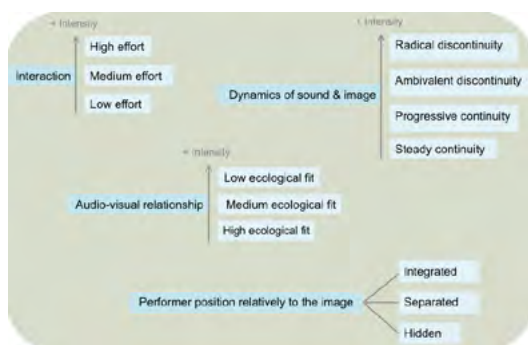

Fig. 1.
Low-level parameters that can be used to analyze interactive performances and installations, non-narrative films and animations, sound art and music, as well as audio-visual performances

3 HIGH-LEVEL PARAMETERS

This section introduces two high-level parameters, which can be used to assess how any low-level variable informs the meaning of the work, and how the product informs the feeling of presence.

3.1 Semantics

We can speak of semantics with respect to causes and concepts, but attention dynamics have intrinsic semantics as well – every experience has meaning, including when we focus on perceptual motion itself. The mental representation of the work as a whole is what we call endogenous continuity (see section 2.3).

The notion of endogenous continuity is useful to expand Jeff Pressing's semantic characterization of sounds [1997], so as to embrace the visual and the audio-visual domain. He distinguished *expressive, informational* and *environmental* sounds, stressing that these typologies are normally overlapped. Expressive sounds would include all kinds of music and song. Examples of informational sounds would be speech, alarms, and sonified data. Examples of environmental sounds would include animal calls, wind sounds, and the noises of machinery. We adapt these semantic typologies as follows:

- **Informational Semantics** prompt causal percepts, shifting attention to a meaning.
- **Expressive Semantics** means that the focus of attention is upon a central target.
- **Environmental Semantics** means a focus upon a context or environment.

These three semantic dimensions can be quantified independently from each other. We can quantify the informational dimension by assessing our conclusiveness about a cause or meaning. It might be useful to look at semiotics, where researchers distinguish three types of relation between signifier and meaning. With icons, the signifier contains something of the meaning (e.g. male & female figures on restroom doors). With indexes, the meaning derives from previous experience (e.g. smoke indicates fire). With symbols, the relation between signifier and meaning is arbitrary, so to say; it depends on convention (e.g. words). These semiotic notions respect to informational semantics, but not to expressive and environmental semantics, which also exist when the informational load is very little. Regardless of concepts and interpretations, the more attention focuses upon a specific target, the less it spreads through the environment, and vice-versa.

When analyzing a creative work, one should consider the semantics of interaction, sound, image, audio-visual relationship and physical setup. One can assess the semantics of each element, and estimate their relative weight in the global meaning of any particular work.

We consider interaction in terms of cognitive effort. Predictable, clearly perceivable interface behaviors provide a large amount of information about how the system should be interacted with; that is why the interaction is effortless. A system that does not depend much on real-time control

is effortless as well, but the informational load of the interaction is little, as the audience does not perceive to which extent the performer is influencing the system output. Furthermore, perceiving effort implies interpreting causes and meanings. Also, perceived effort tends to attract attention, supporting expressive semantics.

Sounds and images have informational load whenever they evoke something beyond themselves. Symbolic systems such as programming code might provide a large amount of information if one understands the code, and very little if one does not. In other cases, the informational dimension can support the expressive or the environmental dimension. For example, a piano recording leads us to imagine a piano and a pianist, and the recording of singing birds evokes a natural environment. Semantic categorization might be less obvious in narrative film. Leo Braudy established a useful distinction between what he called "open film" and "closed film":"In the closed film there is no escape from the logic of actions and events, while in the open film characters may well walk off the frame to some section of the world the camera specifically does *not* define" [1977]. A wide shot provides all the information necessary to interpret the image, leading attention to focus within the limits of the frame. Conversely, a close-up does not describe a scene, leading imagination to build what is not seen within the frame. Thus, a wide-shot of a landscape can be expressive, and a close-up of a person or object can be environmental.

The dynamics of sound and image have their own semantics, be it expressive or environmental. To parameterize dynamics we use the taxonomy of continuities and discontinuities. Steady and progressive continuities create a sense of environment because they fulfill expectations, requiring little cognitive processing; attention can draw to the context/ environment. Ambivalent discontinuities also leave attention under individual control, but they entail more pathos; they attract more attention, reinforcing the expressive dimension of the work. Radical discontinuities make expressive semantics very strong: they prompt automatic attention, monopolizing conscious awareness. Driven by primary instincts, cognitive processing focuses on detecting, perceiving and responding to events in good time.

The informational load of the audio-visual relationship is proportional to its ecological fit, which determines the strength of perceptual binding. This is because binding is informed by concepts of causation [Kubovy and Schutz 2010], which have informational load by definition. Low audio-visual fit provides little information about causation; that is why binding is weak. Medium fit conveys a sense of causation, but the informational load is not very high because one does not understand the base cause-effect relationships. Finally, a high level of audio-visual fit provides a large amount of information; that also explains why it requires little cognitive effort.

The semantics of the physical setup are also important in performances and interactive installations. The central position of a performer, a sound source placed next to him or a spotlight over him will have the effect of directing attention to a central target, conveying expressive semantics.

Conversely, the distribution of sound and light sources in space will emphasize the environment.

A performer's position relative to a moving image is equally influential. A hidden performer confounds the extent to which the work is created in real-time, which means a decrease in informational semantics. With a separated arrangement, the semantics of the physical setup are more expressive when the focus is on the performer, and more environmental when the focus is upon the spatial relation between his body and the moving image. Finally, the semantics of an integrated arrangement can be expressive or environmental. They are environmental whenever the visual output functions like a stage scene, and expressive whenever the physical scale of the work equals that of the human body.

Sometimes, the semantics of the system output are highly expressive due to sonic and visual discontinuities, while the integrated arrangement brings an environmental quality to the work. An example is in Ikeda's *Superimposition*[24]; the performers are in front of a large visual projection, surrounded by multiple video monitors. Another example is in a performance by Metamkine at the Lausanne Underground Film & Music Festival (2012).[25] The performers sit in front of a large projection, using a Super 8 camera, color filters and various devices to create a multitude of light effects and noises. In contrast with these two audio-visual performances, Guy Sherwin created a series of silent performance-films where the integrated arrangement creates expressive semantics, and the image exhibits continuity. In these works, called *Man with Mirror* (1976-2011)[26], he forged a kind of exquisite corps by exploring the relation between his physical body and the changing angles of its reflection on a live-manipulated mirror.

3.2   Performative arena

To complete our parametric model we need a final parameter: one that summarizes how the different semantic dimensions of a creative work intertwine so as to shape the performative arena. This high-level parameter would hardly provide a useful means of analysis if it were considered alone; it would be too subjective. Its role is rather to complement the other parameters, and facilitate the disambiguation of certain aspects. In addition, it can provide cues about elements that have no direct representation, such as the placement of speakers and the lighting.

We distinguish three types of performative arena, which are not mutually exclusive:

- **Local arena** means a focus upon the performer. Expressive semantics are dominant.
- **Distributed arena** means a focus upon the environment. Environmental semantics are dominant.
- **Extended arena** means a subjective sense of presence beyond the physical performance space. It requires perceptual cues, which imply informational semantics.

24. https://vimeo.com/49873167

25. https://www.youtube.com/watch?v=xWwVvICGeR4

26. https://vimeo.com/31609396

27. https://vimeo.cotm/31609396

28. https://vimeo.com/242720751

29. https://www.youtube.com/
watch?v=9mg1weOHWSs

30. http://tarikbarri.nl/projects/versum

31. http://daily.redbullmusicacademy.
com/2017/05/alvin-lucier-music-for-
solo-performer

The local arena relates to Ciciliani's notion of "centripetal" performance tendencies, where the focus is upon the performer [2015]. An unequivocal example is when a sound source is placed next to a musician. Another good example is a video from Sherwin's *Man with Mirror,*[27] where he integrates his physical image and its reflection on a mirror by using a light projector in a dark room, without any light reflections on the wall. The expressive scale of the work is reinforced with informational load, as the interaction is clearly perceivable. Furthermore, a large visual projection can convey the local arena as well. It happens when the performer is separated from the image and the image shows their interaction with the system. An example is in an expanded cinema performance by Arnont Nogyao (2017),[28] where the visual projection shows his interaction with the modified surface of a vinyl record.

The distributed arena relates to Ciciliani's "centrifugal" performance tendencies, where the focus is upon space and context [2015]. A clear example is in Niblock's performances, which create immersive environments. His drone music is played through multiple loudspeakers distributed in space, and the eyes are directed to large visual projections rather than to the performer, who sits in darkness.

The models of Birnbaum [2005] and Ciciliani [2015] consider the relation between the physical and the psychological space of a creative work, but they do not address how the work might expand one's presence beyond the physical space. Shifting spatial presence beyond the physical space requires perceptual cues, derived from the informational load of the sound, the image, or the audio-visual relationship. An example of the extended arena is in Vasulka's performance dedicated to Nam June Paik,[29] where the imagery recalls real-life situations: trees shaking in the wind, Vasulka playing the violin, Michel Waisvisz playing The Hands. Another example is in *Versum,*[30] the musical 3D environment created by Tarik Barri.

The local and the distributed arena are often combined. In Lucier's *Music for Solo Performer,*[31] attention is driven to the performer, who sits still in a central position, using his brain interface to activate percussion instruments. At the same time, attention is driven to the environment, because the instruments and the loudspeakers are distributed in the room, amongst the audience. Another, imaginary example would be a performer carrying a TV monitor displaying footage from the surrounding environment. Indeed, the distributed arena does not require a distributed physical setup. Furthermore, a creative work can combine all three types of arena. For example, Mick Grierson created a 3D composition and improvisation system that uses adaptive algorithms; it looks and behaves like first-person computer games [2007]. We can speak of a local arena because the performer is visible and the interaction design draws attention to their skills. We can speak of a distributed arena because the system takes advantage of a multichannel audio system. And we can speak of an extended arena because the system looks and behaves like a first-person video game – it extends the sense of presence to the digital world.

As a conclusion, we should note that the three types of arena reflect the three semantic typologies but the inverse is not necessarily true. For example, in Steina Vasulka's performance there is no local arena because the performer is not visible, yet the visual discontinuities and sonic deviations create expressive semantics. Similarly, there can be environmental semantics without the arena being distributed or extended; a sonic mass of drones emitted through a single loudspeaker illustrates the idea. Also, the semantics of an audio-visual performance can be informational without the arena being extended, as happens in Magnusson's performances with the Threnoscope system. Indeed, the parameters in our model are intended to complement each other.

Fig. 2.
Variables that influence spatial presence. The blue-shade items have direct parameterization. The blue-font items can be inferred from the high-level parameters - semantics and performative arena.

4   FURTHER RESEARCH

The combination of low-level and high level parameters in the same analysis model offers several advantages. The model provides methods to analyze each variable in particular, and methods to interpret their relationships.
The methods are straightforward when considered individually, but their combination is complex. This makes it useful to represent all parameters graphically, with a set of axes. In practice, the integrated parametric visualization model has been very useful to the analysis of different performance languages, as well as to the development of an audio-visual instrument. Detailed information can be found at http://research.gold.ac.uk/19431/ and http://adrianasa.planetaclix.pt/research/practiceOverview.htm.

BIBLIOGRAPHY

Birnbaum, D., Fiebrink, R., Malloch, J. and Wanderley, M. M.
"Towards a Dimension Space for Musical Devices". In *Proceedings of the International Conference on New Interfaces for Musical Expression,* Vancouver, BC, Canada (2005).

Braudy, Leo
The World in a Frame. Garden City: Anchor Books, 1977.

Boltz, Merilyn
"The Cognitive Processing of Film and Musical Soundtracks". *In Memory & Cognition 32* (2004): 1194-1205.

Brown, S.W. and Boltz, M
"Attentional Processes in Time Perception: Effects of Mental Workload and Event Structure". In *Journal of Experimental Psychology: Human Perception and Performance,* 28 (2002): 600-615.

**Ciciliani, Marko
and Mojzysz, Zenon**
"Evaluating a Method for the Analysis of Performance Practices in Electronic Music". In *Proceedings of the Second International Conference for Live Interfaces* (2015), ISBN 978-989-746-060-9

**Collins, N., McLean, A.
Rohrhuber, J. & Ward, A.,**
"Live Coding Techniques for Laptop Performance". In *Organised Sound* 8 no. 3 (2003): 321-30.

Emmerson, Simon
Living Electronic Music, Hampshire, U.K.: Ashgate Publishing Limited, 2007.

Knudsen, E. I.
"Fundamental Components of Attention". In *Annual Review of Neuroscience 30, no.1* (2007), 57-78.

Moritz, William
The Dream of Color Music, And Machines That Made it Possible. Animation World Magazine, 1 April 1997. http://www.awn.com/mag/issue2.1/articles/moritz2.1.html

Pressing, Jeff
"Some Perspectives on Performed Sound and Music in Virtual Environments". In *Presence: Teleoperators and Virtual Environbments 6, no. 4* (1997).

Sá, Adriana
"How an Audio-Visual Instrument Can Foster the Sonic Experience". In *Live Visuals,* eds. L.Aceti, S. Gibson, S. M. Arisona, O. Sahin, Leonardo Almanac vol. 19 No. 3, MIT Press (January 2013): 284-305. http://journals.gold.ac.uk/index.php/lea/article/view/84

Sá, Adriana,Caramiaux, Baptiste and Tanaka, Atau
"The Fungible Audio-Visual Mapping and its Experience". In *Journal Of Science And Technology Of The Arts vol. 6 No.1* (2014): 85-96. http://artes.ucp.pt/citarj/article/download/131/80

Sacau, A., Laarni, J., Hartmann, T.
"Influence of Individual Factors on Presence". *In Computers in Human Behavior* 24, 5, Elsevier Science Publishers B. V. (2008), 2255-2273

Schutz, M., and Kubovy, M.
"Causality and Cross-Modal Integration". In *Journal of Experimental Psychology: Human Perception and Performance 35, 6* (2009), 1791–1810

Chion, Michel
Audio-Vision: Sound on Screen. Trans. C. Gorbman, New York, Columbia University Press, 1994.

Draper, J. V., Kaber, D. B., & Usher, J. M.
"Telepresence". In *Human Factors 40,* no. 3 (1998): 354–375.

Grierson, Mick
"Noisescape: A 3d Audiovisual Multi-User Composition Environment". In *The World as Virtual Environment*, ed. J. Birringer, T. Dumke, K. Nicolai, Dresden: Trans-Media-Akademie Hellerau (2007): 160-168.

Kubovy, M. and Schutz, M.
"Audio-Visual Objects". In *Review of Philosophy and Psychology 1, 1,* Springer Science and Business Media B.V. (2010), 41–61.

Pressing, Jeff
"Cognitive Complexity and the Structure of Musical Patterns". In Generative *Processes in Music - The Psychology of Performance, Improvisation, and Composition,* ed. by J. Sloboda, New York, Oxford University Press (1987), 129-178.

Sá, A., Ryan, J., McPherson, A., Tanaka, A., Magnusson, T., Van der Heide, E., Grierson, M., Carvalhais, M., McLean, A.
"Live Interfaces: Problematising Foundations". In *Proceedings of the Second International Conference for Live Interfaces* (2015), ISBN 978-989-746-060-9

Sá, Adriana
"Repurposing Video Game Software for Musical Expression: a perceptual approach". In *Proceedings of New Interfaces for Musical Expression 2014* (London, 2014): 331-334. http://nime.org/proceedings/2014/nime2014_343.pdf

Sá, Adriana
"Designing Musical Expression". In *XCoAX 2017 Proceedings.* http://2017.xcoax.org/pdf/xcoax2017-Sa.pdf

Schubert, T.
"A New Conception of Spatial Presence: Once again, with Feeling". In *Communication Theory 19,* no. 2 (2009):161–187

Voegelin, Salome
Listening to Noise and Silence: Towards A Philosophy Of Sound Art. Continuum, 2010.